MASSACHUSETTS INSTITUTE OF TECHNOLOGY

ARTIFICIAL INTELLIGENCE LABORATORY

A. I. Memo No. 530 May 1979

Using Enhanced Spherical Images for Object Representation

David A. Smith

Abstract. The processes involved in vision, manipulation, and spatial reasoning depend greatly on the particular representation of three-dimensional objects used. A novel representation, based on concepts of differential geometry, is explored. Special attention is given to properties of the enhanced spherical image model, reconstruction of objects from their representation, and recognition of similarity with prototypes. Difficulties associated with representing smooth and non-convex bodies are also discussed.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Office of Naval Research under Office of Naval Research contract N00014-77-C-0389.

Introduction and background

Dealing with any task involving physical objects in the three-dimensional world requires a suitable representation. Computer scientists have used and proposed a variety of these representations since the first attempt to computerize such an undertaking. The earliest models, deployed by pioneers in the field, maintained an object as an ordered set of endpoints (i.e. corners) or pairs of points (i.e. lines or edges). Though these had the attractive features of making the object representation explicit in a geometric sense and of facilitating straightforward display of the particular structural instance and spatial orientation, their downfall resulted from the fact that data manipulations of most any sort (and thus transformations and manipulations, which effect spatiotemporal changes, and prototype similarity recognition of unknown objects) were indeed computationally cumbersome.

As early as 1963, Roberts [18] was attacking the problem of machine perception of three-dimensional scenes involving polyhedra. The broad goals of this pioneering work included the construction of line drawings from a pictorial source, the recognition of the three-dimensional objects present in such drawings, and the redisplay of these forms after performing hidden line removal. In the computer, polyhedra were represented as sets of line segments, each corresponding to a physical edge; the desired manipulations were then effected by means of the now classic matrix transformations applied to vectors in a homogeneous coordinate system. A thorough treatment of these techniques is given by Newman and Sproull [14]. The computational demands that this

method imposes even for simple transformations (e.g. translation, rotation, and scaling) in addition to the cumbersome nature of attempts at object recognition have encouraged researchers to pursue representations which are more conducive to the domain of use.

By 1971, Binford [3] had developed an innovative alternative representation for curved objects. His scheme was to model objects in terms of "generalized cylinders", i.e. all physical objects in a system would be transformed in some way to a conglomeration of cylinders of varying size, shape, and orientation. He thereby introduced the notion of representing all solid bodies as assimilations of objects taken from a small set of primitive solids. In Binford's case, a cylinder, the single primitive, consisted of a space curve, or axis, and a circular cross section function on this axis. A simple object could be modeled by choosing a single space curve and an associated cross section function so as to yield a good fit with respect to the shape and dimensions of the surface, and complex objects could then be a composite of generalized cylinders. Of course, a sacrifice was made in exchange for this simplified model -- exact representation became impossible; and it was highly probable that the mappings of objects into the representation were not unique, making recognition difficult.

Shortly thereafter, two independent, assembly-oriented systems evolved. That of Braid [4] was based on a collection of six primitive solids -- cuboids, wedges, tetrahedra, cylinders, sectors, and fillets. Voelcker [19] based his representation system on cuboid and cylinder primitives. The idea of using generalized cylinders or cones to approximate parts of objects after segmenting them into appropriate pieces has been further

developed by Marr and Nishihara [12]. A discussion of various other approaches to three-dimensional geometric modeling in a visual system (such as by statistical world models, procedural knowledge, or semantic knowledge) is presented by Baumgart [2].

Each of these representations has several drawbacks, the major ones are difficulty of computational manipulation, difficulty of object recognition, and an incompatibility between the representational form of the object and the object description available as "input". In this light, a representation which facilitates ease of manipulation and recognition and which is a "natural" form for the object seems to be a most desirable attainment. It is these qualities which the ***enhanced spherical image*** **(ESI)** representation maintains.

In simplest terms, an ESI model for a convex object consists of a set of vectors. An individual vector's direction component represents the direction of a surface normal of the object, and the vector's magnitude signifies the object's surface area corresponding to the particular normal direction. Alternatively, considering each vector of the set to be of unit length and locating them all at a common point of application,[1] the locus of points consisting of their endpoints lies on the surface of a unit sphere. Such a mapping was developed by Gauss and is known as a spherical image. [6] A scalar value, which represents the surface area corresponding to the normal vector, may then be tagged to each point. This model of an object has been dubbed an ***enhanced***

---

1. From differential geometry, a vector in 3-space consists of two points in 3-space: its vector part $\underline{V}$ and its point of application $\underline{P}$. The vector $\underline{V}_p$ may then be pictured as the arrow from the point $\underline{P}$ to the point $\underline{V}+\underline{P}$. [15]

spherical image.

Representational advantages

With the success of several computerized techniques for determining object shapes from photographic sources -- such as photometric stereo [21], which uses two or more images of the same scene taken with varying directions of incident illumination, and so-called "shape from shading" algorithms [7, 8] which determine orientations by augmenting gradient space schemes with image intensity information -- digitized representational data may be produced. To organize this information into an easily manipulated and thus useful form, it needs to be encoded into a complementary object representation. The shape determination process actually resolves the object surface normal at each point in the scene. With this knowledge, the surface area of the various regions which compose an object may be easily computed. (Note that the normal vectors' directions determined for patches of a single planar region are identical to within a reasonable degree of certainty.) A single vector may then be used to represent the direction in which a flat surface is facing. Pairing this vector with the region's area and doing this for all of an object's regions (i.e. faces) will yield an ESI representation.

The question that immediately arises concerns the case in which some faces of an object are not visible to the shape determining process. Certainly the observable regions may be accurately represented. Under such circumstances, a useful property of ESI's may then be employed, specifically that referred to as the *center of mass property.*

(See page 28 in Appendix II.) With this, an object's description may be checked for completeness ("wholeness"); and if a region is missing, the appropriate direction and area needed to fill in the hole may be calculated and appended to the object's representation.[1]

Once an object has been represented in this fashion, the commonplace object manipulations become trivial to execute. It is assumed here that the representation is maintained as a set of vectors, all of which have a common point of application (a local origin in Cartesian space) and whose vector parts are specified in spherical coordinates relative to this local origin. Translations do not effect ESI representations. Rotations about the point of application are effected by adding the appropriate directional offset to the two components of each vector's direction specification or, equivalently, maintaining a base vector which would be added to each vector on any reference to any vector's direction (i.e. rotating the sphere).[2] Rotations about an arbitrary axis in the Cartesian space may then be effected by a rotation about the local origin followed by the appropriate translation. To do scaling, the area associated with each ESI point would be multiplied by the square of the desired scale factor -- which could also be maintained in a base vector. These methods are significantly less computationally demanding than would be the case for a point-edge representation of an object.

---

1. With respect to reconstruction, however, it is inevitable that an object's faces which are adjacent to the contrived face will be somewhat deformed in shape if more than a single planar surface was missing from the description.
2. It may, however, be simpler to represent the vector parts of the ESI points in Cartesian coordinates so that a 3 x 3 orthonormal matrix may be used to maintain the rotation information.

Reconstruction

An ESI representation of an object is not in itself "viewable" or "displayable". In fact, it is not clear that there is a straightforward method for obtaining a point-edge representation of an object from its ESI. What is needed is a method to locate the planes which compose the object faces at the proper distance from a center point. The vector direction specifies the direction for the normal vector of the plane, while the intersection of adjacent planes must define a region of the proper area. Techniques have already been developed for generating an object from a set of intersecting planes. [22]

A recursive algorithm for reconstruction has been suggested but not implemented. (See page 35 in Appendix II.) The proposed proof of correctness has been shown to have an error: a pair of ESI points may be arbitrarily close and yet not represent adjacent faces of a polyhedron. One should note, however, that this does not prove the algorithm itself to be incorrect.

An iterative, reconstructive algorithm has also been envisioned though there remains considerable uncertainty about its convergence. One may conceive of sliding planes in and out from a common point of application along the ray coinciding with their normal vectors, trying to adjust each plane's position so as to produce polyhedral faces with the required areas. Specific techniques have not been investigated.

One advantage of the ESI representation is a simple method of hidden line elimination. Once an object has been reconstructed, its surfaces and edges have been determined. Then only those faces which correspond to ESI points that lie in the hemisphere of the unit sphere which is observable by the viewer would be made manifest in the actual display. Conveniently, these faces are exactly those that the viewer would see if he were examining the three-dimensional object, while the points from the opposite hemisphere are the portions of the object which the viewer would be unable to see from his current observation position.

Recognition

Another intended advantage of modeling objects by their spherical images is that one can easily recognize shapes by comparing their representations. For polyhedra, the task may be pictured as that of trying to rotate two concentric spheres of unit radius such that the dots on their surfaces match up (i.e. for each dot on one sphere there is a dot on the other sphere in the same position). Since the dots may be of various "intensities" or "masses" (recall that the area is encoded in some fashion), these values also need to agree for there to be a match.

The problem involves three degrees of freedom (two for the axis direction plus one for the rotation position about the axis) and is equivalent to holding one sphere in a fixed orientation while matches are attempted by rotating the other sphere about many different axes. One may envision various clever schemes for implementing this

trial-and-error approach to the matching process, such as only trying those axes which pass through ESI points, rotating the sphere to match one pair of points, and then comparing the other points. To determine, however, that two representations of N points each are not the same type of object, such a scheme would try on the order of N^2 different positions of the sphere and would require comparing N points for each.

If, on the other hand, a canonical orientation could be found for each sphere, then, by first aligning them according to their characteristic orientations, only one relative position of the spheres need be tested and only N pairs of points need be checked to determine a match. To develop such a promising technique, it is convenient to regard the ESI points as point masses, the representation of a polyhedron thus becoming a system of point masses. Obviously, the problem's three degrees of freedom are eliminated if two axial directions on a sphere are fixed. Moments of inertia for physical bodies are determined with respect to a given axis, so one naturally thinks of the axes of minimum and maximum moment of inertia -- with the added restriction that the axes must pass through the center of mass of the body. The problem of finding a characteristic orientation for an ESI may then be viewed as that of discovering the axes about which a system of point masses would have the minimum moment of inertia and the maximum moment of inertia.

This task is known in Lagrangian dynamics as that of determining the ***principal moments of inertia.*** [13] The techniques as applied to the problem discussed above are developed completely in Appendix I. Using those results, it is possible to assign a

unique, characteristic orientation to any ESI representation.[1]

Recognition by principal axes alignment

The first step in recognizing that an object matches a known prototype is to align the two spheres so that their characteristic axes coincide. The next step involves determining whether a true match of the ESI points exists. In the case where the object representations are guaranteed to be exactly correct (e.g. in an artificial or synthesized example), the points must coincide perfectly and have identical masses (surface area values). Matching becomes trivial. In more realistic situations, however, one might expect that a match should still occur even if some of the points are slightly amiss in location or mass. This would be especially true if the unknown object's representation was the result of processing real world data, such as a photograph. The matching algorithm must then incorporate some tolerance, the amount of which could perhaps be dependent on the known uncertainty associated with the shape determination process.

At this point, only a couple of speculative suggestions can be mentioned. First, it can be noted that if an exact match exists, then the alignment of principal axes will pair ESI points without any discrepancy -- provided unique principal axes exist. This

---

1. In general, a single unique orientation may be determined; but for objects which exhibit a high degree of symmetry, two or more equivalent orientation frames may be possible. (Consider, for example, the axes about which a cube has the maximum moment of inertia.) This does not seem to pose any problem to the matching process, because each of these possible orientations is equivalently useful. In simply choosing one, a correct match will still result, since the match is based on shape and size alone -- a cube has no unique "top" side. Yet, a regular tetrahedron is an exception to this general idea. It allows many matches which are equally good based on moments of inertia, but not all of which are genuine.

means that in no case will it be possible to slightly rotate one sphere to produce a superior match. However, in a situation where an exact match is not obtained, it is not certain that alignment of principal axes will produce the "best" or "closest" match. If it is found that a slight rotation in some situations would result in a better match, a technique known as *chamfer matching* seems ideally suited for use in determining the best refinement to make. [1] A pictorial explanation of how chamfer matching works would be to view each ESI point on one sphere as being down in an inverted cone shaped valley. The points on the other sphere wish to roll down into the valley as if by gravity, but the system of points is rigidly connected. The matching finds the most agreeable compromise.

In either case, the remaining issue is that of determining the relative goodness of a match. It is probable that one would wish to recognize an object mostly by its shape, rather than its size. (Should a little cube "match" a big cube?) This means that the difference in location between ESI points is significant; whereas, it is the ratio of masses for each pair of points that one should examine. Potentially appropriate measures would be the root-mean-square of the differences in position between paired points and the standard deviation of the mass ratios of point pairs.

This manner of recognizing objects seems quite promising for objects all of whose features are known. Unfortunately, from a single view (as in a photograph) it is usually the case that parts of the object are unobservable and must be guessed at. The ESI representation itself aids in filling in missing pieces, but any attempt to identify such

an object by the recognition method just described would be hopeless in all but a few lucky cases. At present, the only method envisioned for recognizing such objects would involve preserving the hemisphere of valid points and then trying to apply a similar, though appropriately modified, method using them alone -- determining principal axes, etc. The prototypes against which the match was being done must also be "halved." Of course, one need only consider hemispheres which contain the desired number of points. This technique results in many more matches being tried (on the order of the number of points in the prototype).

In the preceding discussion of this method of recognition, it has been implicitly assumed that the representation of the object is at least not too grossly in error and, specifically, that each represented object face corresponds to exactly one real face of the object being modeled. If, however, the shape determination process reports a face that does not actually exist or if it fails to detect a face that is actually present, then the resulting ESI will have more or fewer points than it should. In such a situation, it seems that the proposed recognition algorithm would fail miserably due to its heavy reliance on matching point-for-point. Yet hypothesizing that the shape determination algorithm would only produce such a mistaken result for relatively small surface areas of the object and assign to them an approximately correct orientation, it might be the case that the characteristic orientation of an erroneous model would differ only slightly from that of the true, accurate model. In such a situation, if the two models were aligned by the orientations, then all but the false ESI point(s) would be closely paired. One would expect the points of the inaccurate model which lie closest to the false point (or location

of the "missing" point) to deviate more from their correct positions than other, more distant, points. Noting such an occurrence during the matching process, the absence or presence of an extra point might potentially be detected, and the correct match could then result. Incorporating a detection scheme of this sort into the recognition process would, however, eliminate part of the screening procedure proposed for determining those representations which should be used in the comparisons. Specifically, if one is seeking to identify an ESI which consists of N points, then in the naive manner one would only examine the prototype models which consist of N points. In dealing with errors that effect the number of presumed faces, it would then be necessary to additionally check the models consisting of N+1, N–1, N+2, N–2,..., N+i, and N–i points, where i is the maximum number of potentially erroneous faces (perhaps a fraction of N, though hopefully a small number).

Representing smooth or non-convex objects

The ESI representation and techniques developed thus far are best suited for application to convex polyhedra. In trying to extend the domain to non-convex objects or smooth-surfaced objects, several difficulties arise. For instance, the ESI for non-convex objects is not unique, i.e. there is not a one-to-one mapping between a non-convex object and its ESI. (See page 37 in Appendix II.) Then too, the spherical image of smooth objects does not consist solely of points, as does that of polyhedra. [6] In fact, an object which has no planar surfaces would have a spherical image that would be distributed over every part of the sphere with varying "density" or "intensity"

providing the shape information. The integral of this density over a patch on the Gaussian sphere equals the total area with surface normal with direction falling within the patch.

Though it may be feasible to represent smooth, convex objects by their true spherical images, no techniques for doing so have been conceived. Instead, the approach has been to employ polyhedral techniques. This necessitates approximating smooth-surfaced objects by polyhedra and resolving the difficulties associated with representing non-convex polyhedra. Expedient methods for approximating complex surfaces by planar faces have been developed. [10] They generally produce a somewhat optimal triangulation of the surface; in other words, the approximating polyhedron for a smooth object would consist of triangular shaped faces. This is quite acceptable for the intended purposes.[1]

Only an inelegant solution has been envisioned to circumvent the ambiguities associated with non-convex polyhedra. This entails using existing algorithms for subdividing the objects into convex components, keeping track of how the pieces "fit back together" somehow. [9] Since each piece is then a convex polyhedron, it may be represented in the normal fashion as an ESI. Reconstruction would involve first reconstructing each component part and then joining the pieces back together in the

---

1. The shape determination processing may already output specifications of the surface normal for many small regions of a surface. This could then be regarded as a many-sided polyhedral approximation to the surface, thus foregoing the triangulation process mentioned above. One should note that it would not be necessary to determine the edges between each small region, though the surface area associated with each would be needed. (In actuality, the little surface patches may not match up exactly at their edges, so they would not be defining a "real" polyhedron.)

proper manner. [9] However, it should be noted that this process would yield a polyhedron. If a smooth object is desired, a smoothing operation would be mandated, though an exact reconstruction of the original object would probably be impossible. Little extra difficulty would result for performing spatial manipulations on such an object; all of the component parts would just be treated as an indivisible set -- operators being applied to a base vector for the complete object.

Recognition of non-convex objects might, however, present more of a difficulty. For instance, one cannot even be confident that the algorithm for subdividing an object into convex components would slice two identical objects in the same way, because in general there is no unique set of slicings to produce convex parts for any given non-convex object. If a division algorithm is developed that will always yield the same components for a given object, the matching process would probably become manageable. In fact, before attempting to match shapes of component pieces, the algorithm could first try to recognize an object by the number of pieces and their spatial relations to each other.

Conclusion

This paper has presented the current state of development for an object representation using enhanced spherical images. The representation of convex polyhedra is well understood, while methods for determining point-edge representations from their ESI forms are less so. Spatial manipulations are a promising advantage of the ESI

model, as is object recognition. The methods for matching two ESI's have been investigated, though some specifics need more attention. Smooth-surfaced and non-convex objects present considerable challenges which must necessarily be overcome for the ESI representation to be practical and useful. A first attempt at handling these difficulties involves approximating the objects by one or more polyhedra, and even this idea offers additional perplexities.

Significant progress has been made towards a complete system of support for the ESI representation of objects, but much work lies ahead if the techniques are to ever be utilized in realistic applications.

Appendix I

Determination of Principal Moments of Inertia

The principal moments of inertia for a system of point masses will be determined here symbolically. [20] Throughout this discussion, the Cartesian coordinates of a point will be expressed as the triple (x_i, y_i, z_i), and the mass will be referred to as m_i.

For any system of point masses, its moments of inertia about the coordinate axes -- X, Y, and Z -- may be computed as follows:

$$I_{xx} = \Sigma\,(y_i^2 + z_i^2)\,m_i$$

$$I_{yy} = \Sigma\,(z_i^2 + x_i^2)\,m_i$$

$$I_{zz} = \Sigma\,(x_i^2 + y_i^2)\,m_i$$

Likewise, its products of inertia may be found:

$$I_{xy} = I_{yx} = \Sigma\,x_i y_i m_i$$

$$I_{xz} = I_{zx} = \Sigma\,x_i z_i m_i$$

$$I_{yz} = I_{zy} = \Sigma\,y_i z_i m_i$$

In general, the inertia matrix (inertia tensor) has the form:

$$\begin{bmatrix} I_{xx} & -I_{xy} & -I_{xz} \\ -I_{yx} & I_{yy} & -I_{yz} \\ -I_{zx} & -I_{zy} & I_{zz} \end{bmatrix}$$

There is usually a unique orientation of axes X-Y-Z at a given origin for which the products of inertia vanish,

$$\begin{bmatrix} I_{xx} & 0 & 0 \\ 0 & I_{yy} & 0 \\ 0 & 0 & I_{zz} \end{bmatrix}$$

This coordinate frame defines the principal axes of inertia, and the corresponding I_{xx}, I_{yy}, and I_{zz} are called the principal moments of inertia. They represent the maximum, minimum, and an intermediate value of the moments of inertia.

The principal moments of inertia are the eigenvalues of the inertia matrix. The solution of the determinant equation

$$\begin{bmatrix} I_{xx}-I & -I_{xy} & -I_{xz} \\ -I_{yx} & I_{yy}-I & -I_{yz} \\ -I_{zx} & -I_{zy} & I_{zz}-I \end{bmatrix} = 0$$

for I yields three roots I_1, I_2, I_3 of the resulting cubic equation. (These are the eigenvalues.)

The direction cosines *l*, *m*, *n* of the principal inertia axes are given by

$$(I_{xx}-I)l - I_{xy}m - I_{xz}n \quad = 0$$
$$-I_{yx}l + (I_{yy}-I)m - I_{yz}n \quad = 0$$
$$-I_{zx}l - I_{zy}m + (I_{zz}-I)n \quad = 0$$

These equations along with $l^2 + m^2 + n^2 = 1$ enable a solution for the direction cosines to be made for each of the three roots (I_1, I_2, I_3) separately.

From the inertia matrix,

$$\begin{aligned}&(I_{xx}-I)(I_{yy}-I)(I_{zz}-I) + (-I_{xy})(-I_{yz})(-I_{zx})\\ &+ (-I_{xz})(-I_{yx})(-I_{zy}) - (-I_{zx})(I_{yy}-I)(-I_{xz})\\ &- (-I_{zy})(-I_{yz})(I_{xx}-I) - (I_{zz}-I)(-I_{yx})(-I_{xy}) = 0\end{aligned}$$

which may be expanded to

$$\begin{aligned}&-I^3\\ &+ (I_{xx} + I_{yy} + I_{zz})I^2\\ &+ (-I_{xx}I_{yy} - I_{xx}I_{zz} - I_{yy}I_{zz} + I_{xz}^2 + I_{yz}^2 + I_{xy}^2)I\\ &+ (I_{xx}I_{yy}I_{zz} - 2I_{xy}I_{xz}I_{yz} - I_{yy}I_{xz}^2 - I_{xx}I_{yz}^2 - I_{zz}I_{xy}^2) \quad = \quad 0\end{aligned}$$

One method for solving such a third order algebraic equation is that of Graeffe's root-squaring method [5], although direct methods for solving cubic equations are also known.[1]

---

1. Intuitively, one would expect that the three roots would necessarily always be real numbers, since they do represent physical moments of inertia and are the result of computation involving only real, positive quantities. This, however, has not been verified by analytical proof.

Once the three principal moments of inertia are determined, they may be used to compute the direction cosines of the corresponding axes, thus giving the characteristic directions for the system of point masses.

$$
\begin{aligned}
(I_{xx}-I)l - I_{xy}m - I_{xz}n &= 0 \\
-I_{xy}l + (I_{yy}-I)m - I_{yz}n &= 0 \\
-I_{xz}l - I_{yz}m + (I_{zz}-I)n &= 0
\end{aligned}
$$

Algebraically solving these equations produces

$$l = (1 + Q^2 + R^2)^{1/2}$$

$$m = \frac{\left(\frac{I_{xx}-I}{-I_{xz}} - \frac{I_{xy}}{I_{yz}}\right)l}{\left(\frac{I_{yy}-I}{-I_{yz}} - \frac{I_{xy}}{I_{xz}}\right)}$$

$$n = \frac{\left(\frac{I_{xx}-I}{-I_{xy}} - \frac{I_{xz}}{I_{yz}}\right)l}{\left(\frac{I_{zz}-I}{-I_{yz}} - \frac{I_{xz}}{I_{xy}}\right)}$$

where

$$Q = \frac{-I_{xx}I_{yz} + I_{yz}I - I_{xy}I_{xz}}{-I_{xz}I_{yy} + I_{xz}I - I_{xy}I_{yz}}$$

$$R = \frac{I_{xx}I_{yz} - I_{yz}I + I_{xy}I_{xz}}{I_{xz}I_{yz} - I_{xy}I + I_{xy}I_{zz}}$$

As a final comment, it should be noted that the computational method presented here for determining principal axes is well-suited for application to point mass systems which do not exhibit a high degree of symmetry. Special treatment would be warranted for cases in which any of the products of inertia vanishes. Conceptually, such a situation might be remedied simply by rotating the system slightly and recomputing the moments and products of inertia.

Appendix II

A Support Paper for The Representation of Three-Dimensional Objects Using Enhanced Spherical Images

David A. Smith

May 1978

Introduction

The representation of three-dimensional objects in the realm of digital computers is a fundamental keystone to several current pursuits in the field of Artificial Intelligence, such as computer vision, spatial reasoning, and manipulation. Straightforward descriptions (e.g., representing an object by its vertices, edges, or surface patches) suffer from the large requirements of data storage and/or of computation time for data processing. For example, to perform a single rotation of a point-represented object about an arbitrary axis using matrix multiplication calculates 384 + 16n products, where n is the number of points [14]. This representation turns out to be quite expensive when methods of object recognition depend on repeatedly rotating prototype objects during some matching process. In this paper, an alternative representation for three-dimensional objects shall be discussed which essentially depicts an object by its surface normals. This will be the Enhanced Spherical Image representation.

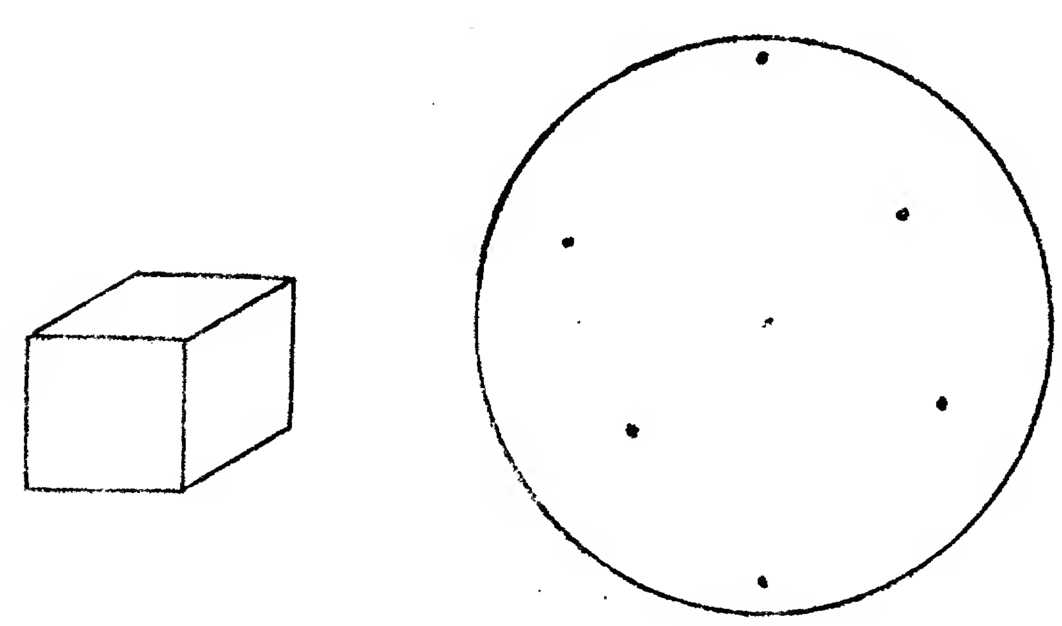
Spherical Images

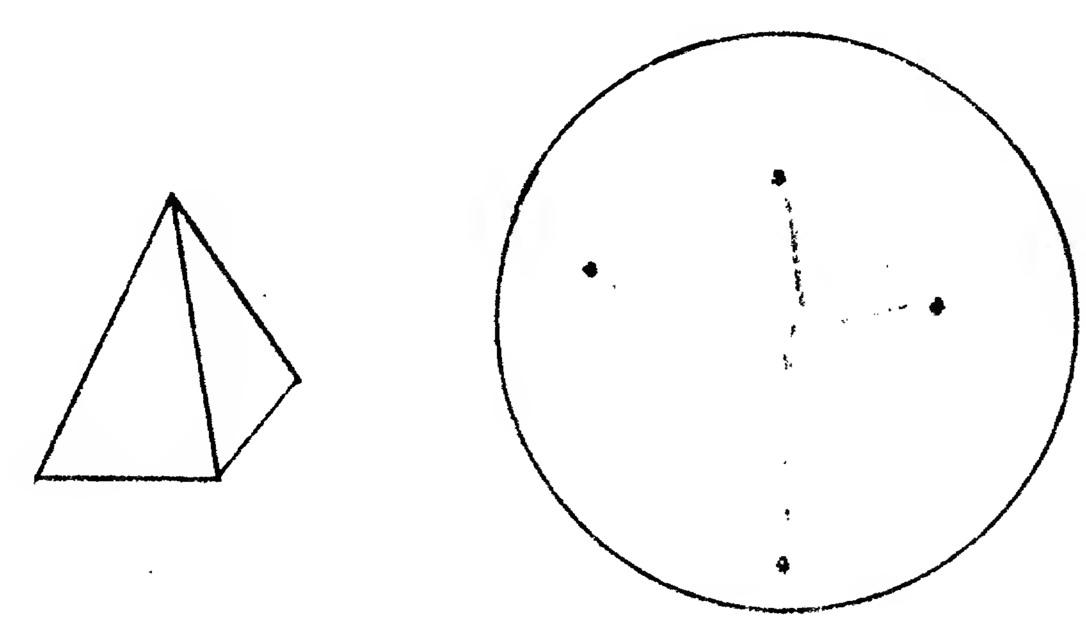
The properties of surfaces in which we are interested include the area, the inclination with respect to some standard direction, and the curvature. The latter may be characterized by two numbers, namely the principal curvatures. These may be thought of as the rate of change of the surface normal directions in each of two perpendicular directions (in the tangent plane). Alternatively, the curvature at a point may be represented by a single number using a method originated by Gauss. This number, of course, will be a function of the principal curvatures and is in fact their product. It is known as the Gaussian

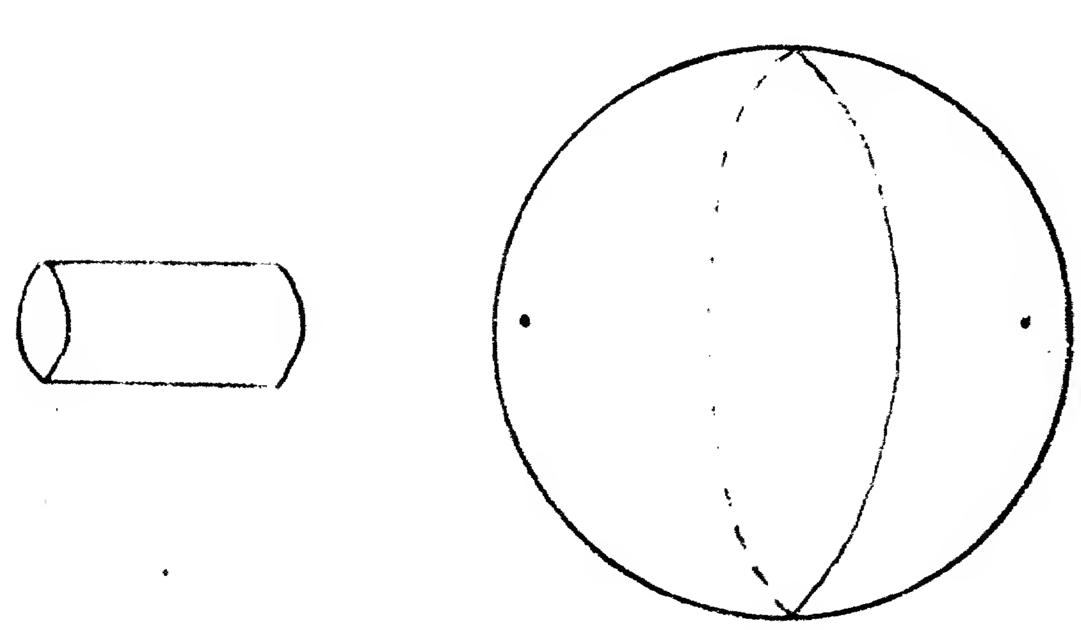
curvature [6, 16].

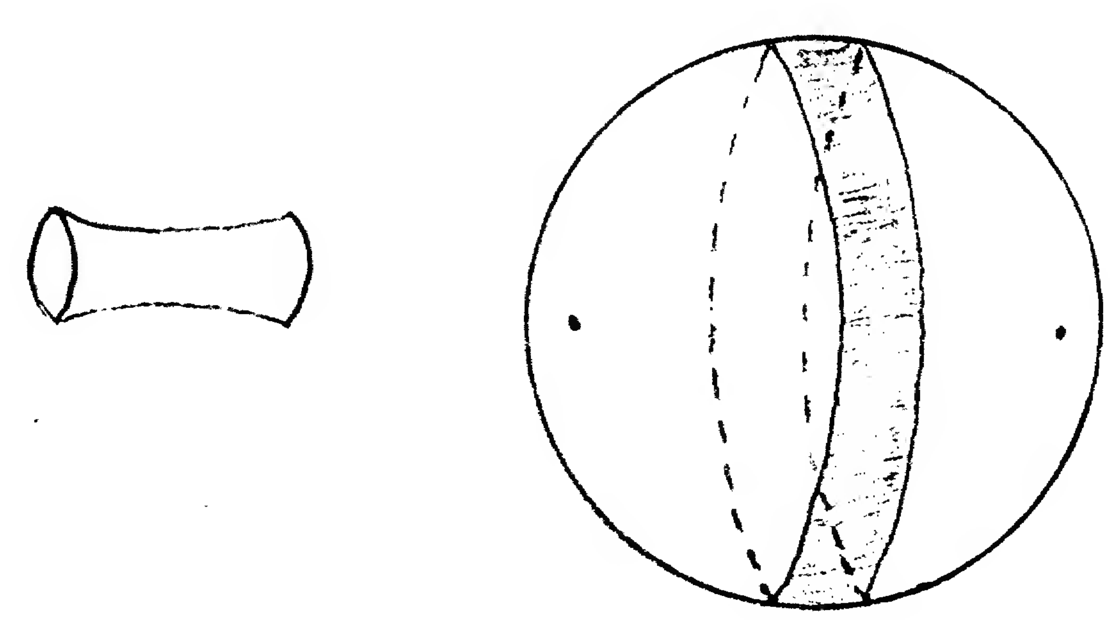
In his studies of differential geometry, Gauss developed a process for mapping a surface onto a sphere. This transformation consists of translating the various outward unit normals of a surface to the center of a unit sphere. The points on the sphere which are at the end points of these vectors are then the spherical representation of the surface, i.e., a *spherical image*. It should be noted that this mapping assigns a definite point on the sphere to every point of the surface.

Throughout the remainder of this paper, the sphere on which a spherical image is formed shall be called the *Gaussian sphere*. Shown below are some simple geometric objects and sketches of their associated spherical images.









Polygons and Polyhedra

Except for the concluding section, this paper shall now mainly deal with the simplified case of representing polyhedra. Therefore, a few of the basic properties of polygons and polyhedra will be presented here to standardize our terminology[11].

A *polygon* is a two-dimensional figure whose sides are straight line segments. A polygon is said to be *convex* if it entirely contains all segments connecting any two of its points. Thus, the *interior angles* (those facing the inside) of a polygon are all less than 180 degrees. A polygon is uniquely determined by its interior angles and the lengths of its sides, though this is not the case if only its angles and its area are specified. (Consider, for example, a 2x3 rectangle versus a 1x6 rectangle.)

A *polyhedron* is a three-dimensional figure whose faces are polygons. Two faces of a polyhedron are *adjacent* if they have a common edge. A polyhedron is *closed* if its faces are convex polygons and all of its edges specify a pair of adjacent faces, i.e., there are no "holes" in the polyhedron. A *convex* polyhedron is one which entirely contains all line segments connecting any two of its points. The angle between any two adjacent faces of a convex polyhedron is less than 180 degrees. Thus, no two faces may have normals in the same direction.

Future references to polygons and polyhedra should be understood to mean convex polygons and convex and closed polyhedra, unless otherwise stated.

In 1897, H. Minkowski proved the following theorem which reveals how polyhedra may be uniquely specified.

> Theorem of Minkowski.
>
> If each face of one convex polyhedron corresponds to a face of another convex polyhedron such that the two faces are equivalent (of equal area) and have parallel outer normals, and conversely, then the two polyhedra are equal and parallel.

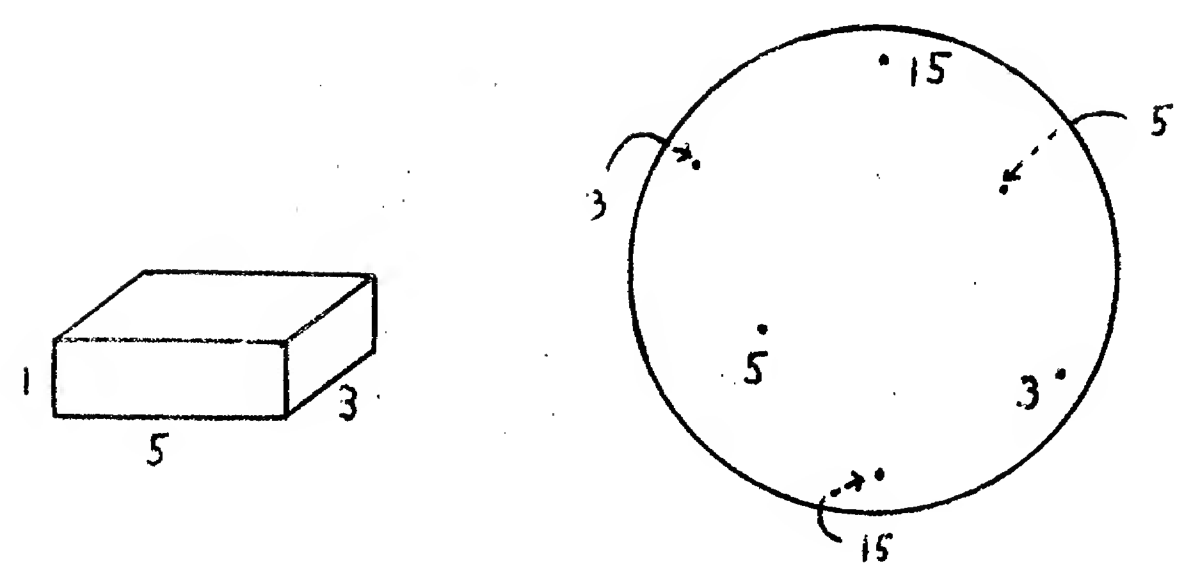
Which is to say that a polyhedron is determined uniquely by the areas and directions of its faces. This concept may now be combined with that of spherical images to develop an efficacious representation for three-dimensional objects.

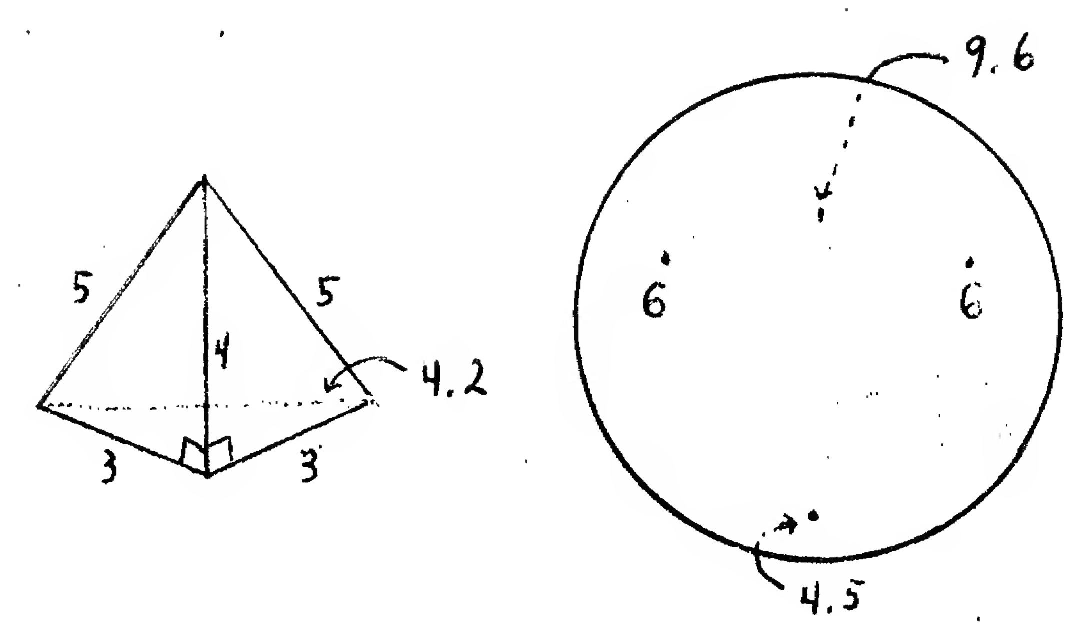
Enhanced Spherical Images

The mapping described above for generating a spherical image does not associate a unique point on the Gaussian sphere with each point of an object, as may easily be verified by considering any body some of whose surface normals are parallel -- one with a planar surface, for example. In the case of polyhedra, the number of normals which map into the same point on the Gaussian sphere is a measure of the surface area of the corresponding face. This area-related data may be preserved by forming an enhanced spherical image (ESI).

> Definition: Enhanced Spherical Image.
>
> An *Enhanced Spherical Image* is a spherical representation in the standard sense (as described above) which has a numeric quantifier, "Intensity", associated with each point of the image. The value of this intensity corresponds to the total surface area of the represented object which has mapped onto this point of the Gaussian sphere.

The following are some examples of convex polyhedra and their associated ESI's.





Uniqueness of an Enhanced Spherical Image

The utility of a depiction for three-dimensional bodies would be questionable if objects were not guaranteed to be represented uniquely. Proof of the following theorem is thus essential to the development of the ESI representation.

Theorem of Uniqueness.

The ESI's of convex polyhedra are unique, i.e., the mapping between a convex polyhedron and its ESI is one-to-one.

Proof.

Part 1: A convex polyhedron maps into one and only one ESI.

This may be shown by observing that the spherical mapping is a composition of one-to-one mappings. We begin by recalling that the outward normals of a surface are necessarily parallel if the surface is planar (as is the case with faces of a polyhedron). Letting each of these normals assume unit length, we may translate (a one-to-one function)

each of them to the center of a unit sphere. Since they all have the same direction, they map into a single point on the sphere's surface (one-to-one). The surface area of one face of a polyhedron is a fixed constant, and once associated with the point in the spherical image, a point in the ESI is formed. This process may be repeated unambiguously (since no two faces have normals in the same direction) for each face of the polyhedron, producing a unique ESI.

Part 2: An ESI of a convex polyhedron maps into one and only one convex polyhedron.

Since no two faces of a convex polyhedron can have parallel normals, a spherical image consisting of n distinct points corresponds to a convex polyhedron of n faces. Thus, an ESI of any convex polyhedron consists of exactly the information which specifies the outward normal and surface area for each face of the polyhedron. By Minkowski's Theorem, we conclude that an ESI of this type represents one and only one distinct polyhedron. Q.E.D.

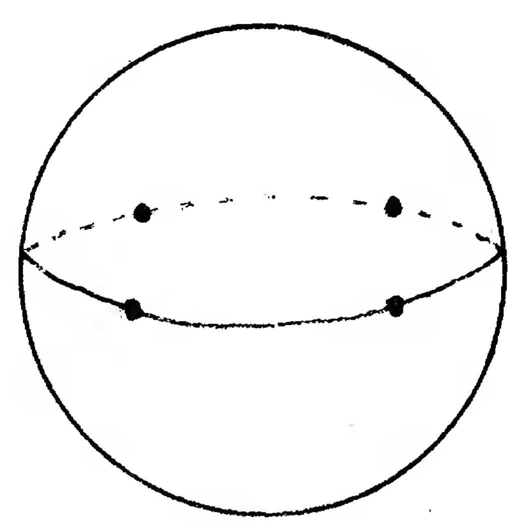
Thus it is shown that the ESI representation may at least be useful in representing convex polyhedra. The reader should observe, however, that certain anomalies arise if the objects are not convex. In fact, for any ESI which corresponds to a real convex polyhedron, there is an infinite set of polyhedron-like objects with holes in their faces which would map into the identical ESI. One can also conceive of infinite sets of ESI's which do not correspond to any real convex polyhedra -- simply permute the intensity or position of one of the points of an ESI which does represent a real polyhedron.[1]

---

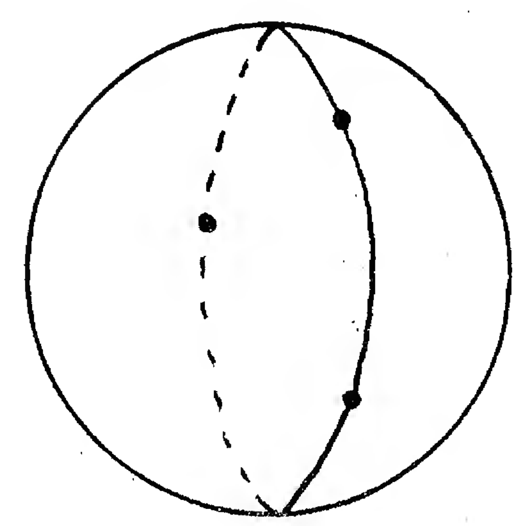
1. If the reader finds this difficult to conceptualize, the Center of Mass Theorem introduced in the next section will provide a more formal justification for this statement.

Another interesting class of objects is that which produces *open* ESI's. An open ESI is one in which all points of the image lie on a single great circle of the Gaussian sphere. A *closed* ESI is one which is not open. For example, figure (A) below could be thought to represent an "open-ended" box, while figure (B) could represent an "open-ended" triangular tube.

A:



B:



For examples of this type, the "length" of the corresponding object is inversely proportional to its cross-sectional area, though neither measure is fixed by the ESI.[1] We proceed now by developing some mechanics towards the utilization of this enhanced representation.

Center of Mass Property

Having observed that not all ESI's, which consist solely of points, actually represent real polyhedra, one can appreciate the need for a simple test to determine the validity of an ESI. As it shall be seen, the test presented here is not only useful in verifying a correctly formed ESI, but it also proves essential to the reconstruction of an object from its enhanced mapping onto the Gaussian sphere.

---

1. It is interesting to note that an open ESI corresponds to an object whose Euler characteristic is not equal to 2, as is the case for all closed objects. Also, a closed ESI corresponds to a bounded convex body, i.e., one which may be enclosed by a sphere of finite radius.

Definition: Center of Mass of an ESI.

The *center of mass*, CM, of an ESI is analogous to the center of mass for a system of point masses. The points of the system are the ESI points, the positions are their coordinates on the Gaussian sphere, and their masses are the intensities of the spherical image points. Thus,

$$CM = \left(\frac{\Sigma m_i x_i}{\Sigma m_i}, \frac{\Sigma m_i y_i}{\Sigma m_i}, \frac{\Sigma m_i z_i}{\Sigma m_i} \right)$$

where (x_i, y_i, z_i) are the Cartesian coordinates of the i^{th} ESI point and m_i is its intensity.

With this definition, we may now state the following valuable theorem.

Center of Mass Theorem.

An ESI corresponds to a real convex object if and only if its center of mass is at the origin of the Gaussian sphere and the ESI is closed.

Proof.

Dividing the Gaussian sphere along one of its great circles, we obtain two hemispheres. Let $\underset{\sim}{N}$ be the outward normal of the sphere which is perpendicular to the plane of the great circle. Then the component of the total area of the convex object which faces in the direction of $\underset{\sim}{N}$ is the sum of

$$m_i * \cos(e_i)$$

over the ESI points on the hemisphere, where m_i is the intensity of the i^{th} point and e_i is the angle between $\underset{\sim}{N}$ and the normal to the sphere at the i^{th} point. If the component of the total area which faces the $-\underset{\sim}{N}$ direction (represented on the opposite hemisphere) is not equal to this sum, then the represented object cannot be closed, and thus cannot be convex. If all pairs of hemispheres do represent equal areas, then all of the point masses are balanced,

and the center of mass must be at the center of the Gaussian sphere. Q.E.D.

Now that the ESI has been shown to represent a convex polyhedron uniquely and that an ESI may be tested for correspondence to a real polyhedron, we may now tackle the problem of reconstructing an object from its ESI representation.

Adjacency, Reduction, and Augmentation

Unfortunately, the information provided by the ESI of a polyhedron does not immediately disclose which of the represented faces are adjacent to each other.[1] This, of course, would have greatly simplified the task of reconstruction. We have not, however, lost all information about the adjacency of the faces, as shall be proven later. Before we invest the effort of doing this, let us examine what we stand to gain. The following definition points us in the right direction.

> Definition: Adjacent points of ESI's.
>
> Two points on a Gaussian sphere are *adjacent* if one of them has no closer neighbors (in terms of the arc length between them) than the other.

Now recalling that the minimum number of faces for a real polyhedron is four (a tetrahedron) and therefore that the ESI of a real polyhedron must have at least four points, we are lead to the following corollary to the Center of Mass Theorem.

---

1. The reader may convince himself of this fact by considering the ESI's corresponding to a cube with one of its corners sliced off. Depending on the size of the removed portion, the new face will have 3, 4, or 5 edges. As long as the direction of the surface normal is kept constant (i.e., all of the slices are in parallel planes), then the positions of the points in the ESI will not change, though the adjacent faces may.

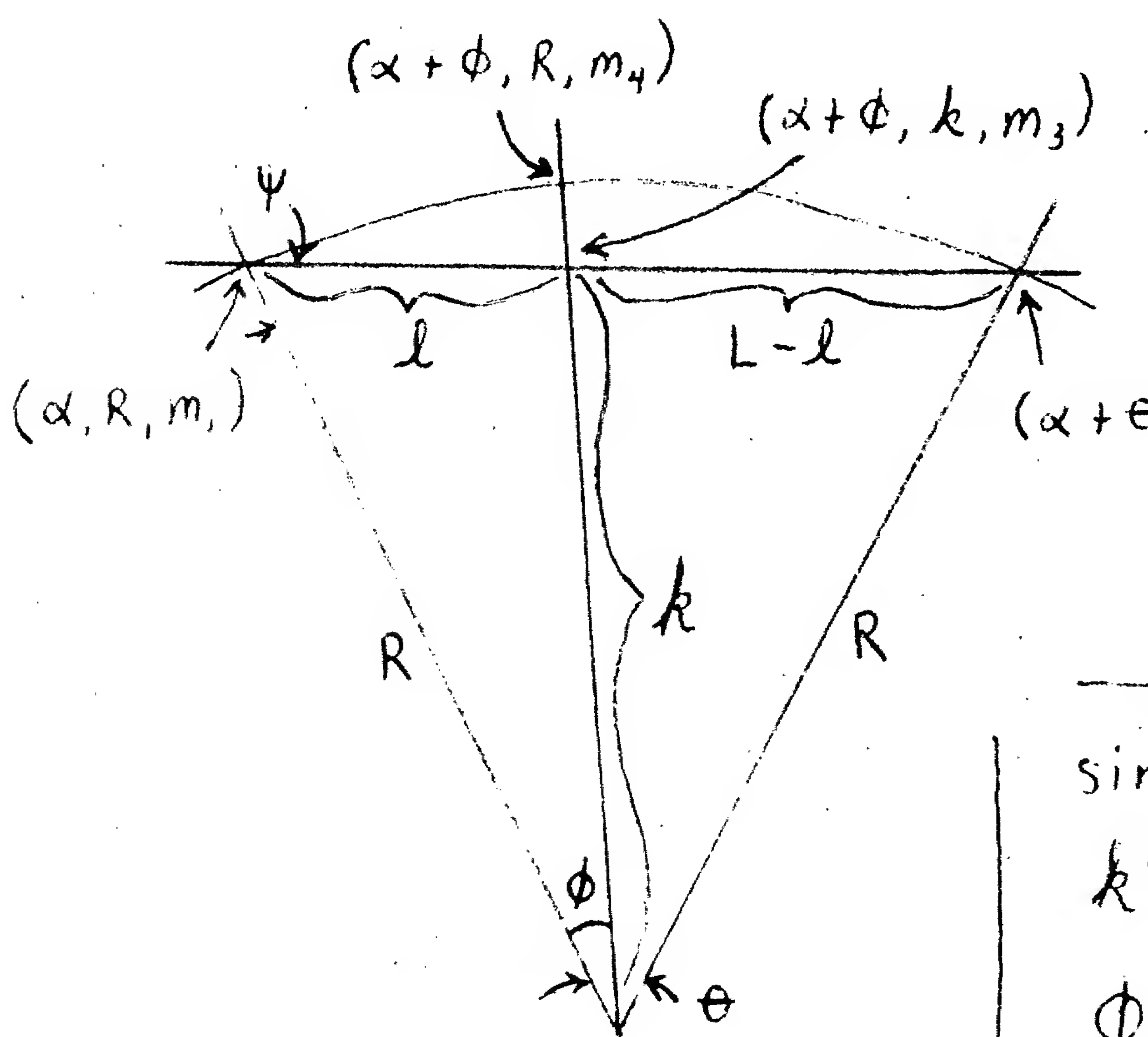
Corollary of Reduction.

> For any ESI which consists of n points (n > 4) and which corresponds to a real convex polyhedron [n-hedron], any pair of adjacent points may be replaced by a single point such that the resulting ESI also corresponds to a real convex polyhedron [(n-1)-hedron]. We shall say that the n-hedron is thus *reduced* to an (n-1)-hedron.

Proof.

The corollary follows immediately from the Center of Mass Theorem if it is always possible to replace two adjacent points by a single point such that the position of the center of mass for the system of points is preserved. We now show that such a point may always be found.

Two points on the surface of a sphere of radius R specify a great circle of the sphere. To preserve the center of mass, the new point must lie along the arc connecting the two adjacent points. The problem is thus transformed into one of two dimensions. Polar coordinates are used here:



Want: ϕ, m_4

$$L = 2R\sin(\theta/2)$$
$$\ell = m_2 L / (m_1 + m_2)$$
$$m_3 = m_1 + m_2$$

$$\sin(\psi) = R\sin(\theta)/L$$
$$k^2 = R^2 + \ell^2 - 2R\ell\cos(\psi)$$
$$\phi = \sin^{-1}\left(\frac{\ell\sin(\psi)}{k}\right)$$

$$m_4 R = m_3 k$$
$$m_4 = \frac{m_3 k}{R}$$

To complement the Corollary of Reduction, one may also give an inverse construction with an inverse proof. This would be the Corollary of Augmentation, and we shall say that an (n-1)-hedron is thus *augmented* to an n-hedron. It may be noticed that this construction is analogous to the augmentation of nets (see p. 56 of[11]).

Theorem of Adjacency

Having the above corollaries, one might now discern the intended plan of attack -- hopefully ending at the ultimate goal of a reconstruction algorithm. The relationship between adjacent points of an ESI and the adjacent faces of the corresponding polyhedron is revealed by the following theorem and shall play a vital role in achieving our goal.

Theorem of Adjacency.

> Adjacent points in an ESI are the mappings of adjacent faces in the corresponding convex polyhedron, i.e., the faces have a common edge.

Proof.

Assume that a pair of adjacent points corresponds to a pair of faces, i and j, of a polyhedron which are not adjacent. Then there exists at least one face, k, between these two faces, and the direction of k's normal vector must differ from that of i more than the direction of j's normal differs from that of i. However, this is possible only if one of the following is the case:

1. There is only one face between the faces i and j, and it has at least one edge that is not the edge of any other face, or

2. There are at least two faces between the faces i and j, and at least one of the inward facing angles between two adjacent faces is greater than 180 degrees.

In either case, the polyhedron is not convex. Thus, the theorem is proven by contradiction.

Theorems of Adjacent Reduction and Adjacent Augmentation

Combining the Theorem of Adjacency with the Corollaries of Reduction and Augmentation produces the final two tools needed for specifying how to reconstruct a polyhedron from its ESI.

Theorem of Adjacent Reduction.

When an ESI of n points is reduced to n-1 points, say by replacing adjacent points i and j with point k, the set of faces of the (n-1)-hedron which are adjacent to face k are the union of those that were adjacent to the faces i and j in the n-hedron (excluding i and j).

The proof follows from the definition of convex polyhedra and the Corollary of Reduction.

Theorem of Adjacent Augmentation.

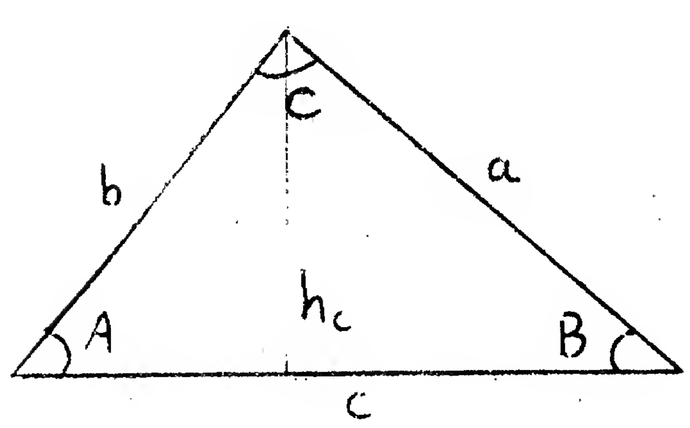
When the faces adjacent to a given face k of an (n-1)-hedron are known, the face k may be augmented into 2 faces i and j, thus forming an n-hedron. The faces adjacent to face i are those whose edges of intersection with k now fall on i's side of the edge between i and j (plus face j), and similarly for j.

The proof follows from the definition of convex polyhedra and the Corollary of Augmentation.

Reconstructing a Tetrahedron

We digress momentarily to examine the "limiting" case, the tetrahedron, which is the polyhedron with the fewest number of faces.

Consider the following property of triangles (which compose the faces of tetrahedra): a triangle is uniquely specified by the length of its three sides. It may also be shown that a triangle is uniquely determined by its area and its three interior angles by deriving the lengths of the sides from this information.



$$K = \text{area}$$
$$= \tfrac{1}{2} h_c c = \tfrac{1}{2} c a \sin(B)$$
$$= \tfrac{1}{2} h_b b = \tfrac{1}{2} b c \sin(A)$$
$$= \tfrac{1}{2} h_a a = \tfrac{1}{2} a b \sin(C)$$

$$c = \frac{2K}{a \sin(B)} = \frac{2K}{b \sin(A)}$$

$$a = \frac{b \sin(A)}{\sin(B)}$$

$$K = \frac{b^2 \sin(A) \sin(C)}{2 \sin(B)}$$

$$b = \sqrt{\frac{2K \sin(B)}{\sin(A)\sin(C)}}$$

Similarly,

$$a = \sqrt{\frac{2K \sin(A)}{\sin(B)\sin(C)}}$$

$$c = \sqrt{\frac{2K \sin(C)}{\sin(A)\sin(B)}}$$

Since the ESI of a tetrahedron provides exactly the information needed to denote the interior angles and area of each face, and since the faces are all adjacent to each other, a tetrahedron may be reconstructed from its ESI as follows:

Tetrahedron Reconstruction Algorithm.

1. Choose one point of the ESI (a face of the tetrahedron).

2. Determine the angles of intersection of the planes represented by the other three points. These give the interior angles of the triangle which is the chosen face.

3. Calculate the lengths of the sides of the chosen triangle.

4. Construct the triangle on the plane specified by the chosen point.

5. Position the other three faces such that their edges of intersection with the chosen face coincide with the sides of the triangle.

Since the positions of all four faces are then fixed, the reconstruction is complete. This special case for tetrahedron will serve as the basis in the restoration of more general polyhedra.

Polyhedra Reconstruction

At last, we have reached our goal. Any polyhedron may be reconstructed from its ESI as follows:

Polyhedra Reconstruction Algorithm.

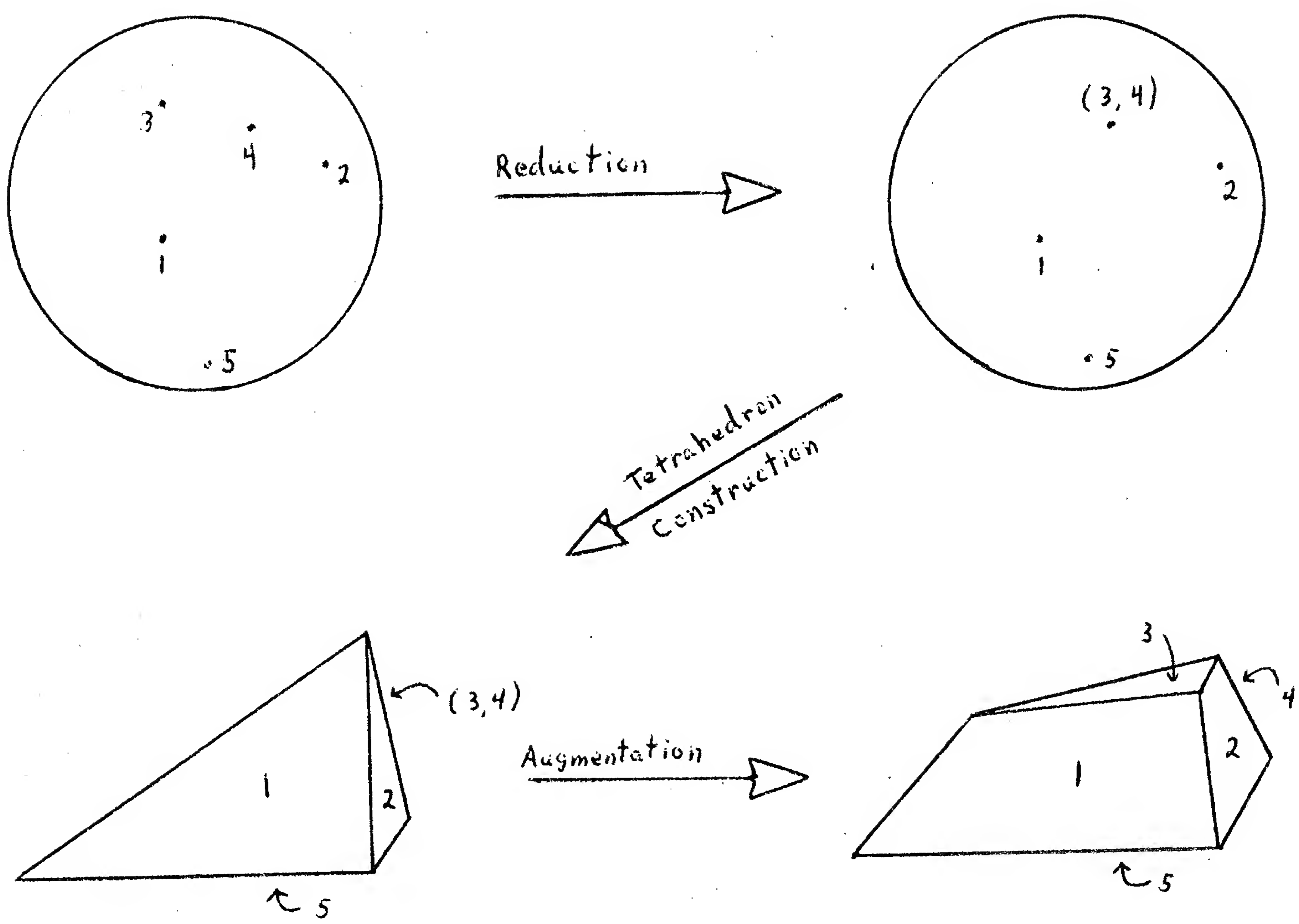
1. Recursively reduce the ESI, "remembering" each pair of adjacent points which get replaced, until it consists of only four points.

2. Reconstruct the tetrahedron which corresponds to this ESI.

3. Successively augment the faces in the reverse order of the reductions using the remembered points to choose the augmentation.

This algorithm may be shown to work correctly by induction on the number of points in the ESI. The series of figures below give a pictorial example of this reconstruction algorithm in

use.



Smooth Convex Objects and Non-convex Objects

For this technique of representing three-dimensional objects to be of practical use, it must be able to model more general objects than simply convex polyhedra. Since the theorems and algorithms developed so far are all inherently dependent on the properties of polyhedra, the initial temptation is therefore to denote smooth convex objects as many-faced polyhedra. This approach suggests examining the limiting situation as the area of each little face then tends to zero. With any luck at all, doing this study will reveal a more general spherical image representation to be utilized in the cases of non-zero

curvature.

Non-convex objects pose a potentially more frustrating problem, since even non-convex polyhedra cannot be represented unambiguously by ESI's. For example, the following two objects have identical mappings onto the Gaussian sphere.

One might consider treating non-convex objects as two or more convex objects with common planar faces. This involves using several Gaussian spheres to represent a single object, and some mechanism for combining reconstructed objects would then be required. Interestingly enough, once we have crossed the threshold of focusing attention on more than a single ESI at any one time, the possibility of representing scenes of objects by a set of "linked" Gaussian spheres then presents itself. At this time, none of these extended schemes have been investigated; but they should certainly be considered in future research involving enhanced spherical images.

Conclusion

An original motive for developing a representation employing surface normals was to facilitate computational manipulation of a three-dimensional object's description. Using the spherical model developed here, rotations about an arbitrary axis are analogous to spinning a sphere -- all points on its surface move at the same time. This is, of course, a result of

using spherical coordinates, where rotation is effected by adding appropriate offsets to each point's altitude and azimuth angles. A more significant advantage of using ESI's would become apparent when object recognition is attempted. Though an algorithm has not yet been developed, the matching process should be greatly simplified. Tasks involving spatial reasoning (as a mechanical manipulator might require), scene analysis, and image reconstruction would all benefit from having an object's surface normals readily available.

It has been demonstrated in this paper that the Enhanced Spherical Image representation is feasible. Suggestions have been offered for future research and possible applications. It is hoped that these further investigations will support the practicality of this representation and that implementation will prove useful.

References

1. Barrow, H. G., J. M. Tenenbaum, R. C. Bolles, and H. C. Wolf, "Parametric Correspondence and Chamfer Matching: Two New Techniques for Image Matching," Proceedings: Image Understanding Workshop, April 1977, pp. 21-27.

2. Baumgart, B. G., "Geometric Modeling for Computer Vision," Stanford University, Memo AIM 249, October 1974.

3. Binford, T. O., "Visual Perception by Computer," I.E.E.E. Conference Systems and Control, Miami, December 1971.

4. Braid, I. C., "The Synthesis of Solids Bounded by Many Faces," Communications of the A.C.M., vol. 18, April 1975.

5. Doherty, R. and E. Keller, Mathematics of Modern Engineering Volume I, John Wiley & Sons, Inc., New York, 1936, pp. 98-128.

6. Hilbert, D. and S. Cohn-Vossen, Geometry and the Imagination, Chelsea Publishing Co., New York, 1952.

7. Horn, B. K. P., "Determining Shape from Shading," Chapter 8, in P. H. Winston (Ed.) The Psychology of Computer Vision, McGraw-Hill Book Company, New York, 1975.

8. Horn, B. K. P., R. J. Woodham, and W. M. Silver, "Determining Shape and Reflectance using Multiple Images," Massachusetts Institute of Technology, A.I. Memo 490, August 1978.

9. Hosaka, M., F. Kimura, and N. Kakishita, "A Unified Method for Processing Polyhedra," IFIP Congress Proceedings, 1974, pp. 49-53.

10. Keppel, E., "Approximating Complex Surfaces by Triangulation of Contour Lines," IBM Joint Research Development, January 1975.

11. Lyusternik, L. A., Convex Figures and Polyhedra, Dover Publishing, Inc., New York, 1963.

12. Marr, D. and H. K. Nishihara, "Representation and Recognition of the Spatial Organization of Three Dimensional Shapes," Massachusetts Institute of Technology, A.I. Memo 416, May 1977.

13. Meriam, J. L., Engineering Mechanics Volume 2 Dynamics, John Wiley & Sons, Inc., New York, 1978.

14. Newman, W. F. and R. F. Sproull, Principles of Interactive Computer Graphics, McGraw-Hill Book Company, New York, 1973.

15. O'Neill, B., Elementary Differential Geometry, Academic Press Inc., New York, 1966.

16. Pogorelov, A. V., Differential Geometry, Noordhoff-Groningen, The Netherlands, 1956.

17. Reddy, D. R. and S. Rubin, "Representation of Three-Dimensional Objects," Department of Computer Science, Carnegie-Mellon University, April 1978, CMU-CS-78-113.

18. Roberts, L. G., "Machine Perception of Three-dimensional Solids," Massachusetts Institute of Technology Lincoln Laboratory, Technical Report 315, May 1963.

19. Voelcker, H. B., and A. A. G. Requicha, "Geometric Modeling of Mechanical Parts and Processes," Computer, December 1977.

20. Wells, D. A., Schaum's Outline of Theory and Problems of Lagrangian Dynamics, Chapter 7, McGraw-Hill Book Company, New York, 1967.

21. Woodham, R. J., "Photometric Stereo," Massachusetts Institute of Technology, A.I. Memo 479, June 1978.

22. Zolnowsky, J. E., "Topics in Computational Geometry," Stanford University, February 1978, STAN-CS-78-659.